AGU100 ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR025500

Tracking Multidecadal Lake Water Dynamics with Landsat Imagery and Topography/Bathymetry

David Weekley[1] and **Xingong Li**[1]

[1]Department of Geography and Atmospheric Science, University of Kansas, Lawrence, KS, USA

Key Points:
- Water shoreline boundaries are an effective means of estimating surface elevation and other dynamics given available elevation data
- Seasonal and long-term trends in lake levels were accurately identified
- The method provides a scalable framework deployable within s allowing rapid assessment of lake dynamics

Supporting Information:
- Supporting Information S1

Correspondence to:
X. Li,
lixi@ku.edu

Citation:
Weekley, D., & Li, X. (2019). Tracking multidecadal lake water dynamics with Landsat imagery and topography/bathymetry. *Water Resources Research*, *55*, 8350–8367. https://doi.org/10.1029/2019WR025500

Received 4 MAY 2019
Accepted 16 SEP 2019
Accepted article online 10 OCT 2019
Published online 1 NOV 2019

Abstract Water resource management is of critical importance due to its close relationship with nearly every industry, field, and lifeform on this planet. The success of future water management will rely upon having detailed data of current and historic water dynamics. This research leverages Google Earth Engine and uses Landsat 5 imagery in conjunction with bathymetry and Shuttle Radar Topography Mission digital elevation model data to analyze long-term lake dynamics (water surface elevation, surface area, volume, volume change, and frequency) for Lake McConaughy in Nebraska, USA. Water surface elevation was estimated by extracting elevation values from underlying bathymetry and digital elevations models using 5,994 different combinations of water indices, water boundaries, and statistics for 100 time periods spanning 1985–2009. Surface elevation calculations were as accurate as 0.768 m root mean square error ($CI_{95\%}$ [0.657, 0.885]). Water volume change calculations found a maximum change of 1.568 km^3 and a minimum total volume of only 23.97% of the maximum reservoir volume. Seasonal and long-term trends were identified, which have major affects regarding regional agriculture, local recreation, and lake water quality. This research fills an existing gap in optical remote sensing-based monitoring of lakes and reservoirs, is more robust and outperforms other commonly used monitoring techniques, increases the number of water bodies available for long-term studies, introduces a scalable framework deployable within Google Earth Engine, and enables data collection of both gauged and ungauged water bodies, which will substantially increase our knowledge and understanding of these critical ecosystems.

1. Introduction

Water is one of the most abundant resources upon Earth and is also one of the most critical to life. While water is massively abundant when considering the Earth's surface as a whole, its distribution is uneven across both time and space (Crétaux et al., 2016) leading to extensive impacts and implications for biology, ecology, economy, and human welfare. Ocean waters aside, global surface freshwater dynamics are especially vital considering that freshwater constitutes just 0.01% of the global water supply, and yet it supports at least 100,000 different species including humans (Dudgeon et al., 2005). Yet, despite its universal importance, our understanding of continental surface water dynamics is limited.

At the most basic level, water dynamics describe where, when, and how much water is present on the landscape. For lakes and reservoirs, water dynamics are observed through water surface elevation, surface area, volume, and volume change measurements. Historically, most water dynamic information has been derived from in situ gauge networks that measure water surface elevation (height or stage) that can be combined with bathymetric survey data (pre- or post-impoundment for reservoirs) to create storage curves relating elevation to surface area and/or volume. Unfortunately, most water bodies lack in situ monitoring, and the data for many others are unavailable due to legal or institutional restrictions (Alsdorf et al., 2007). This problem exists even in developed nations with dense gauge networks. For example, in the state of Kansas, 60% of the population relies upon 80 reservoirs for their primary or back-up drinking water supply (Rahmani et al., 2018), most of which are unmonitored or do not have publicly available data sets.

To alleviate the issues with in situ monitoring, several remote sensing techniques have been developed to estimate water dynamics. As mentioned earlier, water dynamics are typically monitored through changes in water surface elevation, surface area, volume, and volume change. Water surface area is often the simplest metric to estimate as it can often be directly measured using optical sensors, such as the multispectral sensors onboard MODIS (Moderate Resolution Imaging Spectroradiometer) and Landsat or using SAR (synthetic aperture radar) systems. Water surface elevation can also be directly measured, if not available via in situ gauges, through the use of satellite altimeters such as the dual- or single-frequency altimeters

onboard National Aeronautics and Space Administration (NASA)/National Centre for Space Studies' TOPEX/Poseidon satellite ("TOPEX/Poseidon Fact Sheet,", n.d.) or Geoscience Laser Altimetry System onboard NASA's ICESat (Zwally, n.d.). Unlike water surface elevation and surface area, volume and volume change cannot be directly measured. Total volume estimates require knowledge of underlying bathymetry while volume change, at the bare minimum, requires both water surface elevation and surface area measurements on two separate dates to estimate volume change via pyramidal or conical frustum equations (Alsdorf et al., 2007; Crétaux et al., 2016; Gao, 2015).

The aforementioned techniques have been used in numerous studies to analyze water dynamic changes for several water bodies around the globe including, but not limited to, the following examples. Gao et al. (2012) used MODIS and satellite altimetry to study storage variations for 34 global reservoirs via elevation/surface area/volume relationships. Crétaux et al. (2015) used Landsat imagery, satellite altimetry, and the pyramidal frustum formula to investigate regional water dynamics in the Sydarya River region of Central Asia. Moradi et al. (2014) used digital elevation models (DEMs) and subpixel reprocessed MODIS imagery to estimate total water volume and mean lake level for the Caspian Sea. Zhang et al. (2016) used TanDEM-X DEMs to explore bathymetry and storage of unmonitored reservoirs in Brazil. Tseng et al. (2016) used Landsat imagery and DEMs to track water level changes in Lake Mead by estimating subsurface bathymetry from surrounding topography. Cai et al. (2016) used MODIS and area-based water storage estimation models to analyze 15 years of spatiotemporal water storage dynamics for large lakes and reservoirs in the Yangtze River Basin. Additional works in this area include, but are not limited to, studies by Yuan et al. (2017), Wang et al. (2013), Duan and Bastiaanssen (2013), Liang and Yan (2017), El-Shazli and Hoermann (2016), Avisse et al. (2017), and Jiang et al. (2017).

While the methods used in each of the examples above have added substantially to the existing body of knowledge and undoubtedly will continue to do so, each also possesses a distinct set of limitations. Gao (2015) found seasonal effects in MODIS surface area estimates due to subgrid spatial heterogeneity related to mixed pixels along the water shoreline because of the coarse spatial resolution (250 to 500 m) of MODIS. Moradi et al. used subpixel reprocessing to minimize mixed pixels resulting from the coarse resolution of MODIS imagery. Unfortunately, subpixel reprocessing requires increased computation time, and accuracy of water placement within the overall pixel is algorithm dependent (Moradi et al., 2014; Xiong et al., 2018). Both the works of Avisse et al. (2017) and Zhang et al. (2016) are reliant upon water levels being low enough at the time of observation to capture a complete or near-complete bathymetric profile. Unfortunately, most lakes and reservoirs predate modern remote sensing and/or lack the natural variability that would make this technique widely applicable. Tseng et al. (2016), on the other hand, attempt to alleviate the issue by extending the slope of surrounding terrain to estimate bathymetry. While the method performed reasonably well for Lake Mead, DEM resolution dependencies and increasing uncertainty at low-water levels must be considered before applying the technique elsewhere. Techniques such as those employed by Cai et al. (2016) use empirical models, which estimate storage capacities based on regional measurements. Other studies, such as those by Messager et al. (2016) and Heathcote et al. (2015), which also estimated storage and bathymetric features using regional data, found that the techniques worked well at the regional level but cautioned that the results of individual lakes could contain significant error. Satellite altimetry based studies, on the other hand, which estimate water surface elevation by measuring the return time of signals reflected off the surface are limited due to sensor constraints (Solander et al., 2016). Most altimetry missions were designed for ocean or cryosphere monitoring and lack the resolution needed to discern smaller bodies of water without substantial land contamination. Furthermore, the application of altimetry to inland water monitoring is also limited due to orbital characteristics such as large ground-track spacings, which prevent observation of many water bodies (Solander et al., 2016). While altimeter-based calculations are capable of subdecimeter accuracy, those types of results are generally limited to large lakes with favorable shape and conditions (Asadzadeh Jarihani et al., 2013; J. F. Crétaux et al., 2016). Even studies that use elevation/area/volume relationships derived from bathymetric surveys can have issues in cases where surface area is used to estimate water surface elevation or volume. In this case, surface area is represented by a single value and any errors or deviations in that number, such as those caused by local erosional and depositional processes captured in the imagery but not in the bathymetry, directly affect the resulting elevation or volume calculation.

In 2021, the SWOT (Surface Waters and Ocean Topography) mission is set to launch and will feature increased spatial and temporal resolution over its SAR and altimeter predecessors while also offering

approximately 90% coverage of the Earth's surface (Biancamaria et al., 2016). The SWOT mission is expected to offer water surface elevations within 10–25 cm of actual height (depending upon water body size) and surface area measurements within 15% of the actual area over the course of its 3-year mission life (Solander et al., 2016). While the SWOT mission will provide the most detailed and accurate assessment of the world's surface waters to date, its short mission life will, unfortunately, prevent it from providing climate-scale observations. This shortcoming creates a need for better strategies in assessing currently available data sets to analyze beyond the 3-year SWOT window as well as to truly assess the long-term dynamics of these critical water systems.

This paper builds upon previous work by leveraging the Landsat 5 image archive in combination with bathymetric and digital elevation model data within Google Earth Engine (GEE), a cloud-based geospatial processing platform (Gorelick et al., 2016), to assess summer water surface elevation, surface area, and water volume from 1985 to 2009 for Lake McConaughy in Nebraska. Unlike many previous studies that assessed water indices and segmentation thresholds based upon their ability to correctly identify water from non-water, we use water/land interface to estimate water surface elevation. While mixed water/land pixels are of concern in those studies, they are critical to identifying the water-land edge/boundary in our approach. Several common water indices (Normalized Difference Water Index [NDWI], Modified Normalized Difference Water Index [MNDWI], Automated Water Extraction Index [AWEI]) as well as some index combinations (NDWI + MNDWI) were analyzed and evaluated for their ability to predict water surface elevation from composite Landsat imagery using elevation statistics from lake water boundaries across a range of segmentation thresholds. The most accurate combination of variables and their thresholds, determined in comparison to daily gauged water surface elevation measurements, were then used to calculate surface elevation, surface area, volume, volume change, and lake cover frequency for four summer time periods (June to September) each year from 1985 to 2009. Additionally, linear regression was used to assess whether any long-term trends were present for the lake. Finally, this paper compares the results of the proposed technique to those using established techniques. To do so, water surface elevation was estimated using surface area via elevation/surface area relationships developed from the underlying bathymetry and compared to the proposed technique. All in all, this paper comprehensively tests a wide range of water dynamic monitoring parameters, provides recommendations for application to other study areas, and compares the results to established techniques.

2. Study Area and Data

Lake McConaughy, the largest lake and reservoir in the state of Nebraska, is highly variable in both lake extent and volume over time (Central Nebraska Public Power and Irrigation District CNPPID, n.d.; Joeckel & Diffendal, 2004). The lake, located in western Nebraska (Figure 1), was formed following the completion of Kingsley Dam in 1941 as a means of storing irrigation water for what would eventually become the CNPPID. Kingsley Dam is currently the second largest hydraulic fill dam in the world and, in addition to providing irrigation water to the Tri-County region, contributes electricity via a hydroplant completed in 1984 ("Lake History,", 2016).

Beyond the hydroelectric and irrigation benefits provided by the reservoir, Lake McConaughy, or "Big Mac" as it is otherwise known, is also a highly popular recreation destination. Depending upon the water level, the reservoir features white-sand beaches and numerous swimming, water sports, boating, fishing, hunting, and camping opportunities. As mentioned previously, the water-level of the reservoir can vary significantly from season to season or even month to month. This is due to variable inflows from the North Platte River as well as seasonal water requirements, chiefly for irrigation during the summer growing season. Despite being the largest water body in the state of Nebraska, Lake McConaughy is a poor candidate for monitoring via satellite altimetry as it has received limited crossovers from altimeter instruments (Figure 1). Of the available sensors, only RA-2 (Radar Altimeter 2) onboard Envisat in its 2010 to 2012 (end of mission life) orbit routinely crosses the main body of the reservoir ("Pass locator: Aviso+," n.d.). Altimeters, which are non-imaging, profiling instruments, collect information from all the objects within their footprint simultaneously (Sulistioadi et al., 2015) and are typically best suited for water bodies two to three times larger than the sensor signal footprint in the water overpass region to allow adequate sampling (O'Loughlin et al., 2016). Despite that, altimetry has been used for small water bodies with just a single nadir water crossover point; however, these studies were prone to greater error than those with more sampling points (Asadzadeh Jarihani et al., 2013). As a

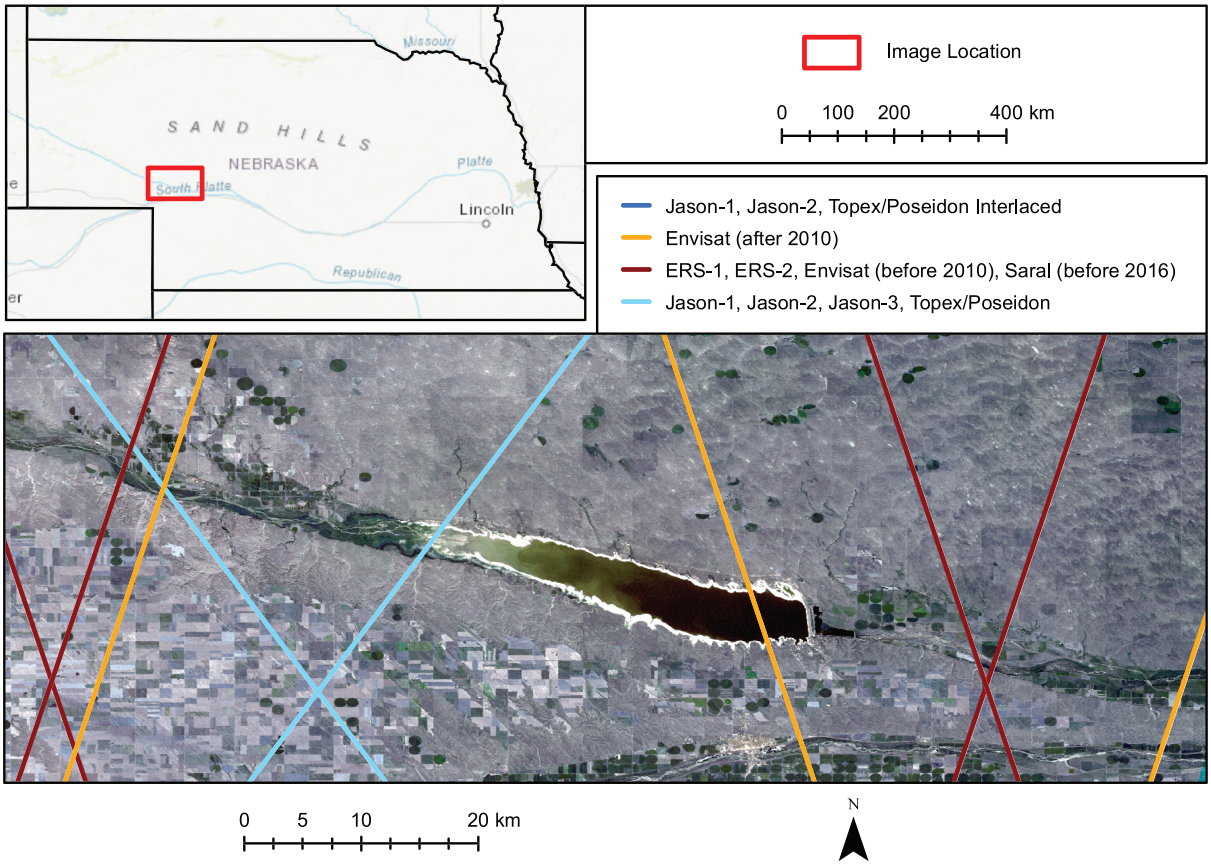

Figure 1. Lake McConaughy and its location in the state of Nebraska along with orbital paths of altimetry missions commonly used for inland water surface elevation monitoring within the vicinity of the reservoir. Of the available sensors with multiple annual crossovers, only Radar Altimeter 2 (orange color) onboard Envisat routinely crosses the main body of the reservoir but contains significant noise. The background image is a natural-color United States Geological Survey Landsat 5 Radar Altimeter 2 Collection 1 Tier 1 Raw Scene from September 2, 2001.

result of these factors, all altimeter measurements near Lake McConaughy, except for Envisat's RA-2, lack the required nadir crossovers to provide water surface elevation estimates for the lake ("RA-2 - Earth Online - ESA," n.d.). Furthermore, Envisat RA-2 only crosses Lake McConaughy from 2010–2012, which is too short a time frame for long-term lake dynamic monitoring.

The primary source of data for this project was GEE, which maintains the entire United States Geological Survey (USGS) Landsat 5 Thematic Mapper (TM) Collection 1 Tier 1 Raw Scenes image archive (1984–2013) with a moderate 30-m spatial resolution as well as the 30-m Shuttle Radar Topography Mission (SRTM) DEM (Farr et al., 2007). The study area, while smaller than an individual Landsat scene, is split between two scene paths (WRS 31/31 and 31/32) and requires imagery from both to cover the entire reservoir. Daily lake elevation gauge measurements were retrieved from the CNPPID in tabular form (H. Rahmann, personal communication, January 9, 2017) and Lake McConaughy bathymetric contours, which were collected as part of a study performed in 2003, were downloaded from the USGS (Kress et al., 2005).

Relatively low water levels at the time of the bathymetry survey necessitated extending the topographic model of Lake McConaughy via merging bathymetry with a supplemental DEM. Our tests indicated slightly better agreement along the land/water interface between the bathymetry and SRTM DEM than the more detailed National Elevation Dataset. To merge the bathymetry with the SRTM DEM, Lake McConaughy bathymetric contours were downloaded, processed, and resampled into a 30-m raster to match the spatial resolution of the SRTM DEM. The overlap area in the SRTM DEM was replaced by the USGS bathymetry data (Figure S1).

Ideally, data for both bathymetry and the surrounding terrain elevation should be collected simultaneously to ensure continuous coverage from underwater to above water. In reality, the time difference between the SRTM mission (February 2000) and the bathymetric survey (Spring 2003) resulted in lake elevation dropping

from a February 2000 average of 992.73 m to a Spring 2003 average of 983.52 m. Due to the higher water level in the SRTM DEM than the bathymetric survey, some stretches of land that were exposed in 2003 but submerged in 2000 were left unsampled and required filling via inverse distance weighting.

3. Methods

Images intersecting the study area (Figure 1) were first retrieved from the USGS Landsat 5 TM Collection 1 Tier 1 Raw Scenes image archive within GEE. Images were then filtered based upon their cloud score (USGS, 2018), sorted into time periods, and processed from digital numbers into top-of-atmosphere composites with per-pixel cloud filtering for each time period. Once the composite images were generated, lake water was identified based on water index images, which were then segmented to create binary images with values of 1 representing water and values of 0 representing land areas. Since the focus of this research is upon one singular body of water, disconnected water bodies were eliminated from the analysis.

Following the removal of disconnected water bodies, a one-pixel radius square kernel was applied over the binary image to extract lake water boundary, which is a ring of pixels around the edge of the lake water body representing the interface between land and water. Water surface elevation was then calculated with statistics (mean, median, or mode) from the bathymetry/elevation values of the pixels within the water boundary.

Once the water surface elevation was determined, surface area, volume, and volume change were then calculated. Surface area calculation was completed with the binary water image by first counting the number of water pixels and then multiplying the count by the area of a pixel. Similarly, volume calculation was done by masking out the nonwater areas from the merged DEM leaving just the pixels covered by water. The elevation of each pixel was then subtracted from the previously calculated water surface elevation to determine the water depth at each pixel. The depth at each pixel was then multiplied by its surface area to determine the water volume of each pixel before being summed to determine total water volume. Finally, water volume change was calculated by subtracting the total water volume of each time period from the successive time period.

While the general processing workflow is outlined above, several key details deserve further explanation including constructing a time series from Landsat 5 images, lake water identification, and water surface elevation estimation.

3.1. Constructing a Time Series from Landsat 5 Images

This research spanned 25 years from 1985–2009 and utilized most of the available images in the Landsat 5 archive. Due to the climate of the study area, the analysis was limited to warm weather months between May 1 and October 31 in order to avoid complications from ice and snow, which would hinder water detection efforts in many early spring, late fall, and winter images. A 3-month temporal window was also used to create four analysis time periods each year: June (May 1–July 31), July (June 1–August 31), August (July 1–September 30), and September (August 1–October 31). While using a time window means that some images might be used in multiple time periods in a year, it was expected that the compositing process would have enough images spread across the entire 3-month window to capture a median value reflective of the central month for that time period (more on the GEE compositing process in subsequent paragraphs). The results of the analysis confirmed this expectation as the values calculated for each time period are quite distinct.

Beyond that concern, the benefit of the temporal window was twofold. First, a larger date range was necessary to ensure that enough images were available for analysis within each time period once images were filtered for cloud cover (discussed in the following section). This is especially important not only due to Landsat's 16-day revisit period but also because the study area crosses more than one Landsat scene. In an early test using single months (June, July, August, and September) as time periods, half of the reservoir was completely devoid of imagery in several months once scenes were filtered by their cloud scores. Second, by overlapping the temporal windows of the time periods, a greater number of analysis periods are available each year, which allows for a more complete picture of water dynamics within the reservoir.

Clouds represent a potentially large source of noise in most remote sensing applications using Landsat images. To address this concern, two cloud removal steps were implemented during the image composition process. First, images with an image cloud score greater than 50% were automatically excluded from the analysis. The image cloud score is a value assigned to each Landsat image that indicates the percentage of a scene

that is covered by cloud. While this metric is useful in identifying cloudy scenes, it does not assess the location or distribution of cloud cover within an image meaning that in some images the object of interest could still be observed and should be included in the analysis even though other areas of the image are obscured by cloud. For this research, tests indicated slightly better results when scenes with image cloud scores greater than or equal to 50% were excluded from the analysis as images with larger values were much more likely to obscure large portions of the reservoir.

As mentioned, the image cloud score alone is a poor measure of the cloud distribution within a Landsat scene. While removing images with cloud scores greater than or equal 50% eliminated many cloudy images, the remaining images could still contain extensive amounts of cloud cover capable of negatively impacting the analysis. To further address this, a per-pixel cloud score generated in GEE during the image composite process was utilized to assess the relative likelihood of a pixel representing cloud cover using a combination of brightness, temperature, and the Normalized Difference Snow Index (GEE, n.d.; Gorelick et al., 2016). While this method is not a robust cloud detection algorithm, it does serve as a simple method for assessing the likelihood that an individual pixel is cloudy. In this research, pixels with relative cloud scores greater than 10 were eliminated leaving the remaining pixels available for image composition. Per-band percentile values at each pixel were then computed from all the remaining pixels to form the final composite image. For this research, all composite images for the time periods were assigned to the 50th percentile value to represent the median central tendency for each time period.

Ideally, per-pixel cloud scores would eliminate the need to filter individual scenes by their image cloud scores. However, early tests that used per-pixel cloud score exclusively to address cloud cover were less accurate overall than using a combination of the two methods where the cloudiest scenes were eliminated from consideration (more in Discussion section).

3.2. Lake Water Identification

In comparison to water identification techniques requiring extensive field observations, training, and validation (such as supervised image classification), segmentation-based techniques, such as the water indices used in this analysis, are very computationally efficient, which are especially useful for analyzing multiple images spanning long time periods. Several water indices were analyzed in this research including NDWI (McFeeters, 1996), MNDWI (Xu, 2006), AWEI (AWEIsh and AWEInsh; Feyisa et al., 2014), and two combined indices formed from NDWI + MNDWI (B1 & B4) and NDWI + MNDWI (B2 & B5). These water indices rely upon the spectral properties of water, most notably its strong absorption in near-infrared and shortwave-infrared wavelengths. Table S1 provides the band designations for Landsat 5 TM while Table S2 provides the formulas for each of the water indices.

Despite computational efficiency, selecting an optimum segmentation threshold for a given analysis can be a difficult process. Atmospheric conditions, water/land composition such as water depth, turbidity and emergent/submerged vegetation, and temporal changes, such as seasons or even time of day, can all influence the optimum segmentation threshold for any single image. This problem compounds itself in cases where multiple images in a long time period are used such as in this study. In addition to this, the different water indices used in this analysis also vary in their requirements for selecting an optimum segmentation threshold. For example, AWEI attempts to produce a stable segmentation threshold (at or near 0 when additional atmospheric corrections are undertaken; Feyisa et al., 2014) while NDWI and MNDWI do not.

While a number of factors affect optimum segmentation threshold selection, we assume that segmentation thresholds could be identified, which minimize measurement error over the entire analysis period rather than attempting to minimize error for a singular time period. In other words, our goal was to identify a threshold or thresholds to accurately estimate water dynamics for the entire image collection in the study period using minimal computational effort rather than find the optimum segmentation threshold for each individual image composite. In order to accomplish this, while searching for the optimum thresholds, the analysis was completed with an increment of 0.01 ranging from −0.35 to + 0.75 for each analysis parameter setting.

Once the water index images were segmented using the selected thresholds, disconnected water bodies were eliminated from binary images using a vector intersect approach, which was performed by seeding a small

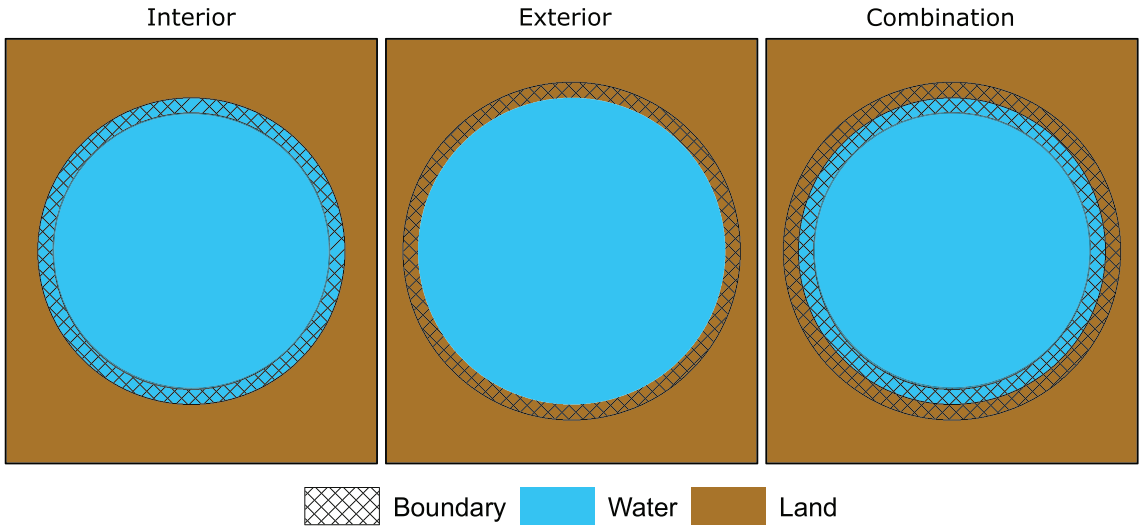

Figure 2. Three potential types of lake water boundaries are possible with a binary water image: interior, exterior, and combination. Water boundaries were identified using a one-pixel radius square kernel to erode and/or dilate the water areas using morphological operations.

polygon (a point or line could also be used) within the main body of the reservoir in an area known to contain water throughout the entire study period. Water areas in the binary water images were vectorized and intersected with the seeded polygon to select the water body of interest, which was then re-rasterized to create a new binary water image (Figure S2).

3.3. Lake Water Identification

We used kernel-based morphological operations to delineate the lake water edge with the goal of using the water/land boundary to determine water surface elevation. Three potential types of boundaries were identified: interior boundary, exterior boundary, and combination boundary (Figure 2). Three boundary types were analyzed because approximating the actual land/water interface on a raster grid is difficult. In reality, shorelines are linear features existing along the land/water interface, but within the raster data, model shorelines inherit the spatial resolution of the raster itself and therefore consist of either land pixels (exterior boundaries), water pixels (interior boundaries), or both (combination boundaries). On a binary water image, when ignoring the effect of mixed pixels, interior boundaries represent the first ring of water pixels while exterior boundaries represent the first ring of land pixels. The combination boundary, on the other hand, uses both land and water pixels to better approximate the land/water interface. Interior boundaries were created by eroding the water pixels on the binary land/water image by one pixel while exterior boundaries were created by dilating the water pixels by one pixel. Combination boundaries were extracted using both methods. While the exact steps varied to some degree depending upon the boundary types used, the boundaries were used to retrieve the elevation from the merged bathymetry/DEM terrain model to capture water surface elevation within one pixel of the water's edge. For a given image, with all other parameters being equal, interior boundaries will produce the lowest water surface elevation estimate, exterior boundaries will produce the highest water surface elevation estimate, and the combination boundary water surface elevation estimate will fall between those of interior and exterior boundaries. With that said, a variety of factors such as local slope, mixed pixels, and water detection accuracy can all impact the ideal water boundary type making a full testing of the available types necessary.

The standard statistics of mean, median, and mode were used to calculate a single water surface elevation from the elevations retrieved from the land/water boundary. Shoreline topography can be very diverse. In some cases, the interface between land and water is a very gentle slope, such as in many beach areas, while other areas can have significant changes over short distances. In areas with low slopes, the depth of water is very shallow making the water surface elevation and the ground elevation essentially the same. Conversely, elevation can differ substantially in areas of steep slope. As such, by using all the values along the shoreline, a single representative value for the water surface elevation can be estimated.

3.4. Validation

Analysis results were assessed by root mean square error (RMSE; Table S3) to determine the best performing parameter combination's ability to estimate surface water elevation for the entire analysis time period. For consistency, and to compare a similar number of data points, the true surface elevations for each time period were calculated using only the dates common to both the daily gauged surface elevations and the Landsat 5 image archive (there are 90+ gauged measurements for each time period versus 12 or fewer Landsat images). The median daily elevation of those common dates was then used as the true surface elevation for each time period. This approach eliminated the influence from events not captured in the Landsat archive, which could have unfairly weighted the gauged surface elevations due to their greater observation density (which increases their chances of capturing events, such as floods, which may or may not be visible within the Landsat image archive due to the 16-day revisit period). RMSE was then calculated using the water surface elevation estimated from image analysis and the true water surface elevation. Additionally, linear regression was used to assess whether any long-term trends were present for Lake McConaughy over the analysis period. Additional error statistics such as mean absolute error, mean bias error, and mean absolute percentage error (MAPE) were calculated for the best performing parameter combination (Table S3). Bootstrap sampling with replacement (1,000 repetitions) was used to calculate each statistic and generate 95% confidence intervals.

Water volume change was evaluated in a similar manner by comparing it to ground truth using error statistics. Additionally, water volume change was also validated against the water volume change calculated using the following pyramidal frustum equation, which is a common technique used in altimeter based studies (J. F. Crétaux et al., 2016):

$$\Delta V = \frac{1}{3}(H_1 - H_0)(A_1 + A_0 + \sqrt{A_1 A_0}) \quad (1)$$

where ΔV is the change in volume between two dates, H_1 and H_0 are the water surface elevations for the two dates, and A_1 and A_0 are the corresponding surface areas for those same dates.

Finally, water surface elevation accuracy was compared to elevation estimates derived from surface area to elevation relationships. First-order (linear), second-order polynomial, and third-order polynomial equations (Duan & Bastiaanssen, 2013) were developed from the USGS provided elevation/surface area/volume table. Water surface areas for each composite image were then used to estimate elevation using the surface area to elevation relationship.

4. Results

All in all, we tested over 5,994 different parameter combinations (water index, segmentation threshold, boundary type, statistical type) to determine the best parameter combination for predicting the water surface elevation, surface area, and volume of Lake McConaughy (Table S4). Each set of parameters was mapped over 100 composite images that were generated from 597 unique Landsat scenes (after cloud score filtering) from a total of 651 scenes available in the study area during the study period. Images were counted a single time even though many were used more than once due to the overlapping time windows.

RMSE was calculated for each parameter combination to allow for accuracy assessment and comparison. Overall, the best performing parameter combination for predicting water surface elevation was NDWI + MNDWI (B1 & B4) with a segmentation threshold of 0.06 using a combination boundary and the mean statistic, which produced an RMSE of 0.768 m ($CI_{95\%}$ [0.657, 0.885]). Table 1 shows the number of combinations by water index with RMSEs better (lower) than some selected values. This table indicates that of the 5,994 total combinations tested, 803 of them produced RMSEs less than 1.0 m (about 13.4%). While this is a low percentage overall, a closer look indicates that the majority of these low RMSEs were concentrated in the NDWI + MNDWI (B1 & B4) and NDWI + MNDWI (B2 & B5) indices, which together account for 519 (64.6%) of the 803 total parameter combinations, which have an RMSE below 1 m. NDWI + MNDWI (B1 & B4) proved to be the most accurate index overall with 10 parameter combinations having RMSEs better than 0.80 m.

Table 1
The Number of Combinations by Water Index Types Which Exceed Selected Water Surface Elevation Root Mean Square Errors

Water indices	Less than 1.25		Less than 1.0		Less than 0.9		Less than 0.8		Total tested combinations
NDWI	212	21.2%	102	10.2%	29	2.9%	0	0.0%	999
MNDWI	228	22.8	128	12.8%	17	1.7%	0	0.0%	999
AWEInsh	10	1.0%	3	0.3%	0	0.0%	0	0.0%	999
AWEIsh	136	13.6%	51	5.1%	4	0.0%	0	0.0%	999
NDWI + MNDWI (B1 & B4)	481	48.1%	245	24.5%	86	8.6%	10	1.0%	999
NDWI + MNDWI (B2 & B5)	488	48.8%	274	27.4%	35	3.5%	0	0.0%	999
Total	1,555	25.9%	803	13.4%	171	2.9%	10	0.2%	5,994

Note. NDWI + MNDWI (B1 & B4), for example, had 245 combinations (24.5%) with root mean square errors better (lower) than 1.0 m and 10 combinations (1.0%) better than 0.8 m. NDWI = Normalized Difference Water Index; MDWI = Modified Normalized Difference Water Index; AWEI, Automated Water Extraction Index.

In addition to the table, the RMSEs for all parameter combinations were compiled into a series of graphs to further highlight and illustrate the impact each parameter has on overall water surface elevation accuracy and allow the comparison between the different water indices, segmentation thresholds, and boundary types (Figure 3). The graphs span the entire range of segmentation thresholds tested (−0.35 to +0.75) but were capped at an RMSE of 2.0 m in order to focus upon the most accurate combinations. The graphs reveal some key patterns and relationships that will be useful in future research. One such pattern is the relationship between boundary types and their lowest RMSE values (or the curve in general). Interior boundaries have their lowest RMSEs at lower segmentation thresholds than exterior boundaries while combination boundaries fall in the middle. For example, the bottom of the mean and median NDWI curves are centered around a segmentation threshold of −0.05 while the exterior boundary is centered around 0.10 with the combination boundary falling in between around 0.0. The graphs also highlight a maximum threshold limit of +0.19 for AWEIsh (Figure 3d) regardless of the boundary or statistic type used as well as the relatively poor performance of AWEInsh in water surface elevation estimation.

Another key observation stems from the type of statistics used in the analysis. Each graph reveals roughly the same pattern where mean has the narrowest RMSE curve, followed by median, and then mode with the widest RMSE curve. Table S5, which shows the number of segmentation thresholds broken down by water index and boundary/statistic type with RMSEs less than 1.0 m, further confirms this observation. Mode, while generally being more stable across a range of thresholds, often performs at a lower accuracy than mean or median. To further illustrate, there are 3 median combinations, 14 mean combinations, and 0 mode combinations with an RMSE better than 0.80 m. Additionally, Table S5 indicates better accuracy with interior (372 submeter combinations) and combination (336 submeter combinations) boundary types compared to exterior boundaries (155 submeter combinations) across the range of tested parameter combinations.

Once the best performing parameter combination (NDWI + MNDWI [B1 & B4], combination boundary, mean statistic, 0.06 segmentation threshold) for predicting water surface elevation was determined additional water estimates for the reservoir were also calculated using those parameters. Figure 4 plots the water surface elevation, surface area, and volume of the reservoir for each month during the study period. Water surface elevation was plotted against the gauged measurements provided by the CNPPID while surface area and volume were plotted against surface area and volumes provided or interpolated, where necessary, in a Lake McConaughy lookup table from the USGS developed from the same bathymetry data used in this study (Kress et al., 2005). Each figure was also fit with a trendline that indicates a moderate negative trend in each of the measurements over the course of the study period. It can also be noted that each year has its own localized trend with the peak elevation, surface area, and volume occurring early summer and then decreasing steadily until the fall.

In spite of the differences, Figure 4 indicates a good relationship between the predicted values and the actual gauge-based values (median values calculated from the daily gauge values and the lookup table). Table 2 provides accuracy assessment statistics for the each of the three estimates.

The estimated water surface elevation, surface area, and volume during the study period peaked in June 1986 at a water surface elevation of 994.54 m. However, the actual peak water surface elevation, surface area, and volume occurred in June 1997 at a water surface elevation of 995.02 m. This error occurs around the

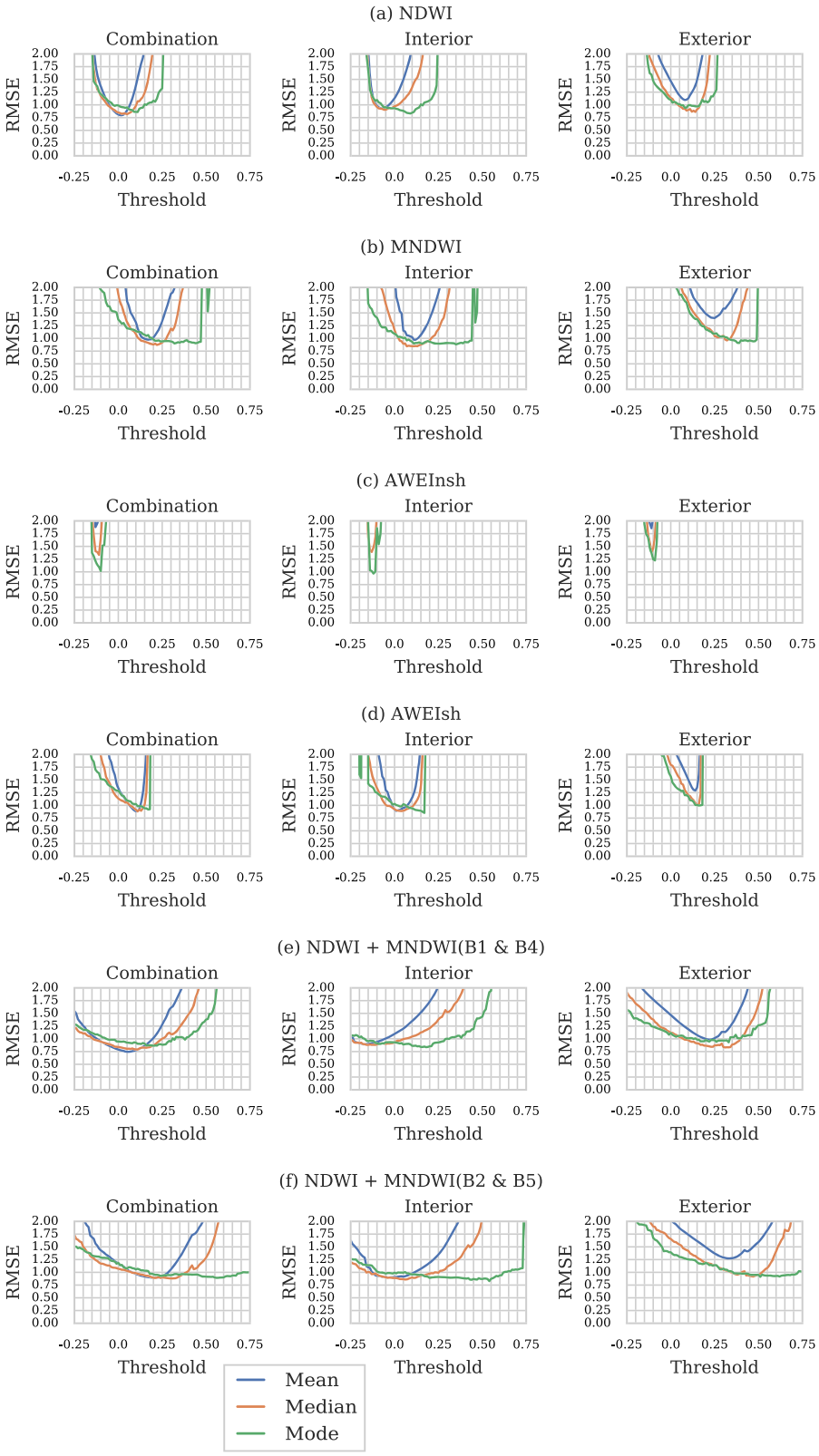

Figure 3. RMSE (root mean square error) curves by water index types. (a) Normalized Difference Water Index (NDWI); (b) Modified Normalized Difference Water Index (MNDWI); (c) Automatic Water Extraction Index – no shadow (AWEInsh); (d) Automatic Water Extraction Index – shadow (AWEIsh); (e) NDWI + MNDWI (B1 & B4); (f) NDWI + MNDWI (B2 & B5).

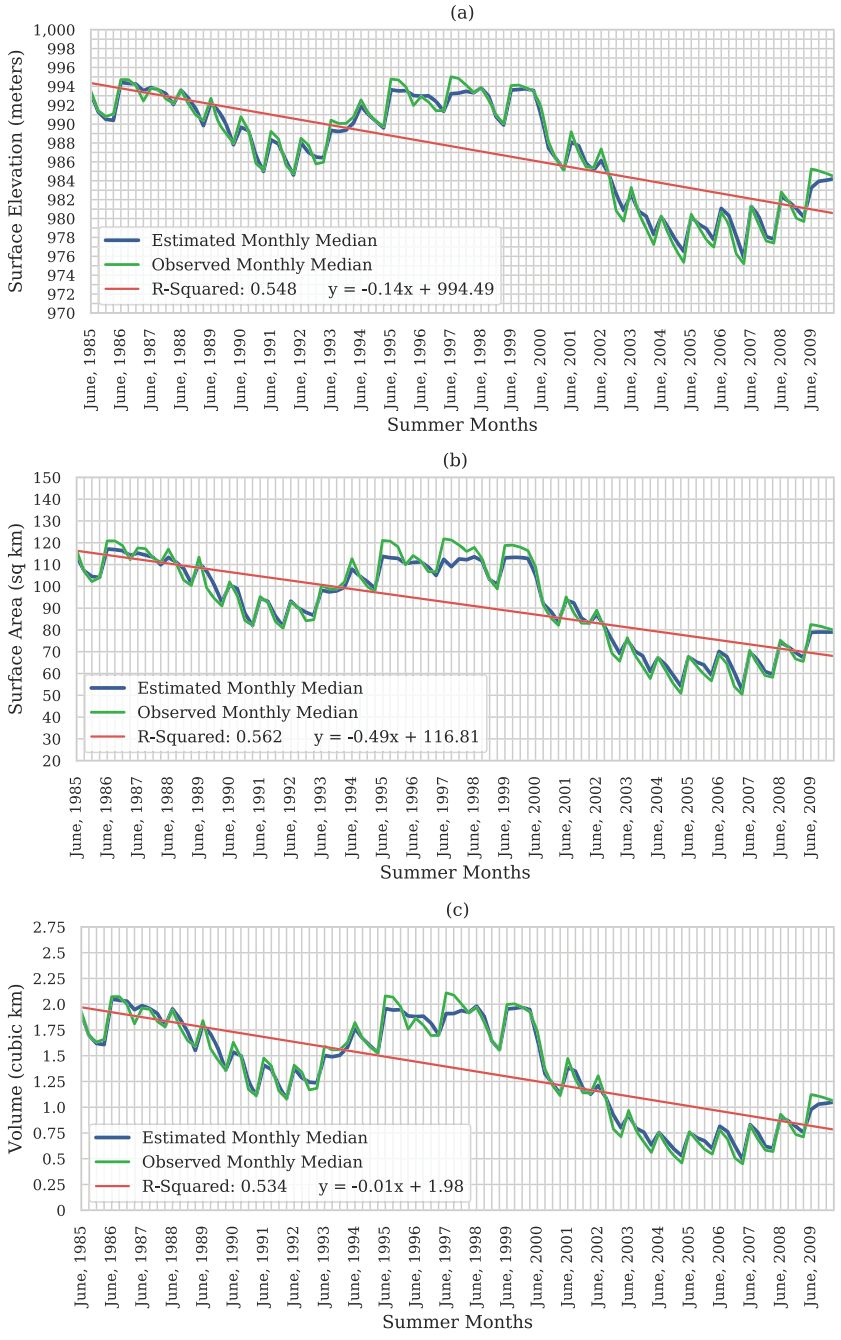

Figure 4. Lake water surface elevation, surface area, and volume over time. (a) Surface elevation; (b) Surface area; (c) Water volume.

SRTM/bathymetry interface and is likely a result of undersampling in those areas as well as estimation bias (see discussion). At the other end of the spectrum, September 2006 had the lowest lake levels for both the estimated (975.91 m) and actual (975.21 m) measurements. Figure 5 provides a longitudinal view of the reservoir to further illustrate the water disparity between the maximum and minimum surface elevations. From Lake McConaughy's peak surface elevation calculated in June 1986 to its minimum in September 2006, the western portion of the reservoir retreated nearly 14.5 km (visible in Figure 5b). In terms of surface cover frequency, 44.04% (~52.25 km^2) of the max reservoir extents was covered by water during all 100 time periods (blue color in Figure S3). Conversely, 0.76 km^2 (0.64%) was covered by detected water just a single time.

Table 2
Accuracy Statistics for Water Surface Elevation, Surface Area, Volume, and Volume Change Using Landsat 5 Image Composites and Merged Bathymetry/ Shuttle Radar Topography Mission Data

Metric	Root meansquare error	Mean absolute error	Mean bias error	Mean absolute percentage error
Surface elevation (m)	0.768	0.601	−0.071	0.06%
	$CI_{95\%}$ [0.657, 0.885]	$CI_{95\%}$ [0.513, 0.698]	$CI_{95\%}$ [−0.223, 0.073]	$CI_{95\%}$ [0.052, 0.071]
Surface area (km^2)	3.527	2.744	−0.009	3.10%
	$CI_{95\%}$ [3.000, 4.085]	$CI_{95\%}$ [2.333, 3.198]	$CI_{95\%}$ [−0.709, 0.710]	$CI_{95\%}$ [2.584, 3.667]
Volume (km^3)	0.069	0.055	−0.012	4.86%
	$CI_{95\%}$ [0.060, 0.077]	$CI_{95\%}$ [0.047, 0.063]	$CI_{95\%}$ [−0.026, 0.002]	$CI_{95\%}$ [3.922, 5.876]
Volume change (km^3)	0.074	0.058	0.00	—
	$CI_{95\%}$ [0.064, 0.084]	CI_{95}% [0.050, 0.067]	$CI_{95\%}$ [−0.014, 0.015]	—
Volume change (km^3)pyramidal frustum	0.074	0.059	0.00	—
	$CI_{95\%}$ [0.065, 0.084]	$CI_{95\%}$ [0.050, 0.068]	$CI_{95\%}$ [−0.014, 0.015]	—

Note. Error statistics for water volume change estimated the pyramidal frustum method are also included for comparison.

Finally, water volume change was calculated. Table 2 shows error statistics for our method, as well as for water volume changes calculated using the pyramidal frustum method (equation (1)) using our estimated surface area and surface elevation. Overall, we calculated a total volume change between the maximum and minimum lake levels of 1.568 km^3 compared to an actual change of 1.659 km^3. At its minimum surface level, the reservoir contained just 23.97% of its maximum volume.

For one final comparison, water surface elevation was estimated for each time period using surface area to elevation relationships built from the underlying bathymetry. Using this method, a maximum accuracy of 0.824 m RMSE ($CI_{95\%}$ [0.706, 0.934]) was achieved using linear regression (first order), a 0.10 segmentation threshold, and the NDWI + MNDWI (B1 & B4) water index combination.

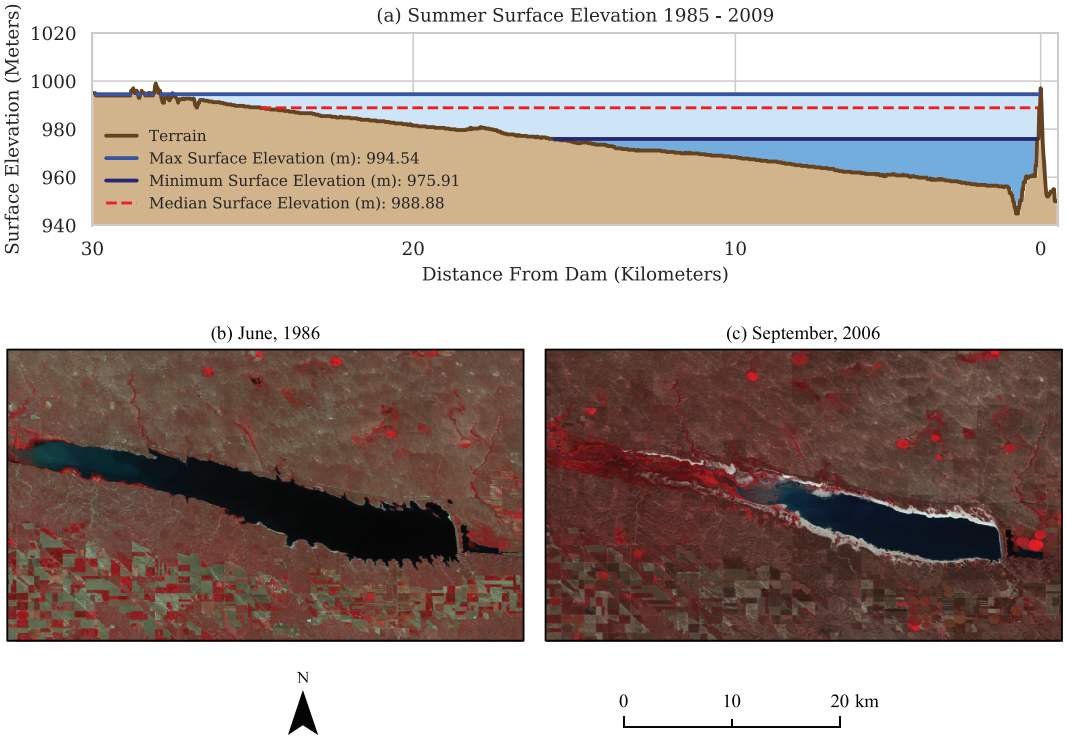

Figure 5. (a) Reservoir maximum, minimum, median summer surface elevations and lake depth along a centralized line bisecting the reservoir longitudinally. It shows the vast change in water quantity between the maximum and minimum reservoir levels. Note that the *y* axis is not in the same scale as the *x* axis. False-color infrared composites of peak reservoir level in June 1986 (b) and minimum reservoir level in September 2006 (c). The western end of the reservoir retreated nearly 14.5 km during that time span.

5. Discussion

5.1. Composite Images and Cloud Cover

During the data exploration stage of this project, we noticed reduced accuracy when using all the available imagery during the study period. We originally thought per-pixel cloud scores would eliminate the need to filter individual images by their overall image cloud scores, but eliminating cloudier images actually improved the results. This is due, in part, to the relationship between the composite images and the ground truth data. The ground truth for each time period is the median gauged surface elevation for each image date. When cloudier images are included in the analysis, the lake is more likely to be obscured resulting in fewer pixels from that date being included in the analysis, which weights the analysis in favor of less cloudy images. The ground truth on the other hand is not subject to cloud cover. In other words, cloudier images result in more data points being used to generate the ground truth value than are being used in the estimation process.

5.2. Segmentation Thresholds, Boundaries, and Statistics

As mentioned previously, the RMSE graphs (Figure 3) reveal several interesting relationships between peak accuracy, segmentation threshold, and boundary types. There appears to be a slight shift in the segmentation thresholds, which produce accurate results among the three different boundary types with interior boundaries typically peaking at a slightly lower threshold than combination or exterior boundaries. Figure 2 illustrates the cause of this difference. Assuming that the same index, segmentation threshold, and statistic are applied, an exterior boundary will be displaced one pixel outward from an interior boundary. In this case, a displacement of one pixel outward for the exterior boundary will likely extract pixels of a higher elevation than would be extracted for the interior boundary. For an exterior boundary and an interior boundary to accurately calculate water surface height, an exterior boundary requires a smaller water surface area to extract the same set of pixels as the interior boundary. Therefore, the peak performance of an exterior boundary would be shifted to slightly higher segmentation threshold than an interior boundary. The segmentation threshold with the best accuracy for combination boundaries typically falls between the best segmentation thresholds for interior and exterior boundaries. As mentioned, one implication of this relationship is that when both boundary types are at the same surface elevation, the interior boundary will therefore provide a larger surface area and a larger lake volume than the exterior boundary, while again, the combination boundary type results would fall in the middle.

Figure 3 also reveals an interesting relationship between segmentation threshold and statistic types. Of the three central tendency statistics, mean is the most susceptible to outlier influence followed by median. Due to varying topography around the reservoir, elevation values captured by the water boundary are sure to include values significantly above or below the actual water surface. Mode, on the other hand, assesses central tendency by determining the most frequently occurring value within the boundary. In this study, while not necessarily accurate overall, mode statistics often produce a significantly wider RMSE curve than mean or median. Mode is able to better leverage the common pixels within the water boundary at each segmentation threshold increment (0.01) to a much greater degree than mean or median resulting in a greater range of acceptable segmentation thresholds that can be used to estimate water surface elevation. However, one caveat of this comes into play when attempting to calculate surface area and volume. For example, using mode, the estimated surface elevation may be very similar using a segmentation threshold of 0.05 or 0.15. However, the 0.05 segmentation threshold still means a greater number of pixels were identified as water when compared to the 0.15 segmentation threshold, which would result in a larger surface area and volume in spite of similar surface elevation. Another potential caveat of using mode statistics could arise when using higher resolution bathymetry/DEMs. With sufficient elevation measurement precision, fewer values would repeat potentially leading to less stable mode estimates.

Additionally, it should be noted that this method may be ill suited for canyon-filled lakes or reservoirs. If the topography near the shoreline is sufficiently steep, then the water surface elevation can diverge significantly from the land elevation above or below water if the DEM/bathymetry does not possess a high enough spatial resolution since this technique does not directly measure the water level itself (such as is done with an altimeter), but rather it assesses the elevation of the shoreline (above or below the water).

5.3. Water Index Performance

Table 1, Table S5, and Figure 3 each highlight a disparity in water index performance across the range of tested parameter combinations. As mentioned, the combined indices, NDWI + MNDWI (B1 & B4) and NDWI + MNDWI (B2 & B5), had better accuracy across a wide range of segmentation thresholds in comparison to MNDWI, NDWI, AWEInsh, and AWEIsh. The improved results of the combined NDWI and MNDWI indices may be due to increased separation between the water and nonwater classes within the image as noted by Acharya et al. (2017) and Lu et al. (2011) in their study of combined water indices. Conversely, the poor performance of AWEInsh and AWEIsh is likely due to a few factors. In terms of this research, the primary cause could be a lack of the additional preprocessing and atmospheric corrections steps undertaken by Feyisha et al. (2014) to improve image quality and produce a stable segmentation threshold at or near 0. In our study, AWEIsh and AWEInsh had the narrowest RMSE curves (Figure 3) and produced the fewest number of submeter surface elevation accuracies of any of the tested water indices. Peak performance for AWEIsh occurred around −0.10 segmentation threshold while AWEIsh generally improved in accuracy until an abrupt decrease in accuracy at +0.19 segmentation threshold. In any case, the peak accuracy performance thresholds for the two AWEI algorithms are significantly different than the stable 0 threshold envisioned by Feyisa et al. (2014). Further research is needed to determine whether the additional preprocessing procedures undertaken by Feyisha et al. would improve the results from the indices.

5.4. Identification of Water Body of Interest

A few potential methods to identify the main water body of Lake McConaughy were examined. In addition to the vector intersect method ultimately used in this study, the cumulative cost (or cost distance) method was also tested and found to be a functional, capable method. The cumulative cost method within GEE does require some prior knowledge about the max dimensions of the area being investigated as it requires a maximum distance parameter to perform the calculation. If the max distance parameter is set too small, the analysis would stop short of capturing the entire water body. Also, if the water body were of sufficient size, using a large max distance could result in memory errors (Gorelick et al., 2016).

The vector intersect method used in this project is a multistep process where all the water bodies within the study area are vectorized. Like the cumulative cost method, the vectorization process can be memory intensive so one potential issue with this method could stem from particularly large water bodies with sufficiently complex shorelines.

5.5. Water Volume Change

The water volume change calculated using our method agrees well with using the pyramidal frustum equation. Despite identical error statistics, our method estimates a maximum volume change of 1.568 km^3 between the maximum and minimum water levels during the study period compared to a pyramidal frustum volume change estimate of 1.541 km^3 and an actual volume change of 1.659 km^3. The decreased accuracy of the pyramidal frustum method may be due to the oversimplification of the lake bottom morphology assumed by the method (Hollister & Milstead, 2010). Also, with an RMSE of 0.074 km^3 ($CI_{95\%}$ [0.064, 0.084]), our method outperformed the 0.6 km^3 RMSE Crétaux et al. (2015) obtained with a combination of Landsat, MODIS, and altimetry, for a reservoir with overlapping water volume ranges with Lake McConaughy. While some of the error differences between our method and that of Crétaux et al. (2015) may be attributable to differences in the lakes themselves, the Crétaux et al. method is also subject to error due to the surface area measurements used to construct the surface area/elevation relationship (coarse MODIS spatial resolution), temporal misalignment of altimeter measurements and surface area measurements from Landsat imagery (passovers occurred on different days), and/or the varying accuracy of the multiple altimeters used in the study. In our study, volume estimation has a MAPE of 4.856%, which is greater than the MAPE of either surface area (3.095%) or surface elevation (0.061%) indicating error propagation throughout the calculations. Increased accuracy for either surface elevation or surface area should improve the accuracies on water volume and water volume change.

5.6. Surface Area to Elevation Relationships

A common method for estimating water surface elevation using optical or SAR imagery is using surface area to elevation relationships developed from the underlying bathymetry. In this scenario, using regression equations developed from the bathymetry, elevation can be estimated using surface area. The results of

this analysis indicate that our method using water boundaries outperforms surface area to elevation relationships (0.768 m, $CI_{95\%}$ [0.657, 0.885]) for our method versus 0.824 m RMSE ($CI_{95\%}$ [0.706, 0.934] for surface area to elevation relationships). We expect that this is the result of using many values (thousands of shoreline elevation values) compared to a single surface area value for estimation. Over long periods of time, erosional and depositional processes may change the underlying bathymetry and cause areas of the lake to expand (increased surface area) or contract (decreased surface area). This change is then captured in the remote sensing imagery but is not accounted for in the bathymetry unless additional surveys are conducted. In this scenario, areas of change have a lower impact on the elevation estimate using shoreline boundaries as they are buoyed by the remaining lake shore compared to surface area to elevation relationships where the lost surface area will immediate result in lower elevation and volume estimates.

5.7. Time Series Implications

Over the course of the study period, two key patterns are visible in Figure 4. First, each year there is a drawdown of the reservoir levels over the course of the summer. The annual summer drawdown occurs because Lake McConaughy was built for irrigation and the CNPPID is required to release water to irrigators upon request regardless of other recreational or environmental needs (Commission, n.d.). In terms of the summer drawdown, Figure 4 reveals some biases in the optimal segmentation threshold results in comparison to the measured ground truth data. In the first low period (1989–1994), surface elevation, surface area, and volume are overestimated early in the year before, generally becoming more accurate in later months. This bias pattern then again repeats itself at the beginning of the second, larger prolonged low period (2000–2009) before the bias pattern reverses once the reservoir falls to lower levels (2003–2008). These biases, as well as some of the error at the highest water levels, likely occur because a single segmentation threshold is being used for the entire analysis period. At high levels, the mixed pixels along the shoreline contain large amounts of vegetation including emergent vegetation, which may interfere with water classification. At low lake levels, the spectral properties of the shoreline change in response to increasing beach area and changing water quality conditions. In essence, the selected model performs better at some water levels than others due to changing spectral properties related to the surrounding nonwater landcover types. Future research, which will include dynamic thresholding and other strategies, should address this issue. As covered in the results, several other parameter combinations also performed well with nine combinations exceeding 0.80 m RMSE and a total of 803 combinations better than 1.0 m RMSE. The total number of combinations with submeter performance illustrates some flexibility in the model selection. However, it is important to acknowledge that each parameter combination has specific water levels in which the model is most accurate and that other water levels will exhibit some bias by either overestimating or underestimating water surface elevation. The strong performance of multiple models and segmentation thresholds further indicates the gains that could be achieved through dynamic thresholding techniques.

Figure 4 also shows a second, long-term drawdown of Lake McConaughy as a result of extended drought periods, which result in decreased inflows and increased irrigation requirements (Joeckel & Diffendal, 2004). Similar to the annual drawdown, but on a larger scale, the long-term drawdown has a direct impact upon recreational activities and ecological habitat. The number and location of available boat docks, as well as the lake's carrying capacity of sportfish, change in response to lake water levels. At the lowest levels, fish mortality is a serious concern as poorer water conditions and potential toxic blue-green algae blooms place increased physical stress on fish and other organisms (Commission, n.d.).

6. Conclusions

The methods developed in this paper have shown great promise for studying long-term lake water dynamics. While the ability to accurately estimate water surface elevation relies upon water indices and selecting an appropriate segmentation threshold, the use of appropriate boundary types and statistical measures can increase the range of acceptable values considerably. Regardless of water indices or boundary type, estimates were least sensitive to changes in segmentation threshold using the mode statistic, followed by median, and finally mean; however, at optimal segmentation thresholds, mean and median provide significantly better accuracy. NDWI + MNDWI (B1 & B4) had the highest accuracy of the tested indices with 10 combinations having an RMSE better than 0.80 me. Overall, 803 of the tested parameter combinations produced RMSEs within 1.0 m of in situ gauge measurements with the lowest RMSE being 0.768 m ($CI_{95\%}$ [0.657, 0.885])

produced from NDWI + MNDWI (B1 & B4) with a combination boundary, mean statistic, and 0.06 segmentation threshold. In general, in applying this method to ungauged water bodies, we would recommend either of the NDWI + MNDWI water indices along with either an interior or combination boundary and the mode statistic. However, if ground truth is available, we would recommend training the model on the available data to further optimize the selected model. Although altimeters are capable of calculating surface elevation at the subdecimeter level, this level of accuracy is generally reserved for significantly larger water bodies with favorable shape, area, and topography. While reliant upon existing bathymetry, the proposed methods in this paper would expand lake dynamic studies to additional lakes for which bathymetry exists but altimeter observations are unavailable. In the state of Kansas, for example, bathymetric surveys have been conducted for approximately 80 lakes and reservoirs, yet relatively, few of these lakes are suitable to altimeter-based studies ("Kansas Lakes and Reservoirs,", 2016). The method could also be used for independent verification of new sensors and missions studying inland water dynamics. Furthermore, this method modestly outperformed surface area to elevation relationships (0.768 m, $CI_{95\%}$ [0.657, 0.885]) for our method versus 0.824 m RMSE ($CI_{95\%}$ [0.706, 0.934] for surface area to elevation relationships).

Additionally, our study successfully revealed the seasonal patterns of Lake McConaughy (highest level in the spring and lowest in the late fall in response to irrigation water diversions during the summer), as well as recognized the moderate long-term trend present over the 25-year time period ($R^2 = 0.547$ for lake surface elevation), which has major implications for not only agriculture in the surrounding area but also recreational activities, fish habitat, and water quality within the lake itself. All in all, the results of lake surface elevation, surface area, and volume were in excellent agreement with ground truth values.

While the methods showed promise in tracking long-term lake dynamics, it is expected that higher-resolution imagery, DEMs, and bathymetry would increase the accuracy of our method as well as improve the ability to calculate volume change between time periods. One potential source of error in this analysis likely occurs along the interface between the SRTM DEM and bathymetry data due to the bathymetry being collected at a lower lake surface elevation than was present during the SRTM mission. Additionally, improved cloud filtering techniques would also boost the final analysis, and dynamic segmentation thresholding techniques could improve accuracy across various surface elevation zones. Finally, this study also assumes a static reservoir bottom for volume calculations. In reality, this reservoir, like many around the world, has been in-filling with sediment over the entirety of its life.

Future research will focus upon improving DEM/bathymetry merging techniques, more robust cloud filtering techniques including FMask (available in the Landsat QA bands), methods of estimating sediment infilling, as well as utilizing higher-resolution and/or temporal resolution imagery products to increase the number of available observations and estimation accuracy. Overall, the long-term monitoring of water dynamics has the potential to improve water resource management as well increase our understanding of temporal changes in water quantity distribution and its impacts upon water-dependent phenomena.

Acknowledgments
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The GEE script used in this analysis is available at https://code.earthengine.google.com/8f3d01ac4acab9d5164c993f1d518270 (Weekley, 2019b), and the resulting data is available on Figshare at https://figshare.com/s/cf6971062945f1954013 (Weekley, 2019a).

References

Acharya, T. D., Subedi, A., Yang, I. T., & Lee, D. H. (2017). Combining water indices for water and background threshold in Landsat image. *Proceedings*, *2*(3), 143. https://doi.org/10.3390/ecsa-4-04902

Alsdorf, D. E., Rodríguez, E., & Lettenmaier, D. P. (2007). Measuring surface water from space. *Reviews of Geophysics*, *45*, RG2002. https://doi.org/10.1029/2006rg000197

Asadzadeh Jarihani, A., Callow, J. N., Johansen, K., & Gouweleeuw, B. (2013). Evaluation of multiple satellite altimetry data for studying inland water bodies and river floods. *Journal of Hydrology*, *505*, 78–90. https://doi.org/10.1016/j.jhydrol.2013.09.010

Avisse, N., Tilmant, A., François Müller, M., & Zhang, H. (2017). Monitoring small reservoirs' storage with satellite remote sensing in inaccessible areas. *Hydrol. Earth Syst. Sci*, *21*(12), 6445–6459. https://doi.org/10.5194/hess-21-6445-2017

Biancamaria, S., Lettenmaier, D. P., & Pavelsky, T. M. (2016, March 27). The SWOT mission and its capabilities for land hydrology. *Surveys in Geophysics. Springer Netherlands.*, *37*(2), 307–337. https://doi.org/10.1007/s10712-015-9346-y

Cai, X., Feng, L., Hou, X., & Chen, X. (2016). Remote sensing of the water storage dynamics of large lakes and reservoirs in the Yangtze River Basin from 2000 to 2014. *Scientific Reports*, *6*(1), 1–9. https://doi.org/10.1038/srep36405

CNPPID. (n.d.). Nebraska Lake Elevations. Retrieved August 6, 2019, from https://www.cnppid.com/news-info/reservoirriver-data/

Commission, N. G. and P. (n.d.). Lake McConaughy Lake Ogallala Master Plan. Retrieved from https://outdoornebraska.gov/wp-content/uploads/2017/03/Lake-McConaughy-Master-Plan_final.pdf

Crétaux, J. F., Abarca-del-Río, R., Bergé-Nguyen, M., Arsen, A., Drolon, V., Clos, G., & Maisongrande, P. (2016). Lake volume monitoring from space. *Surveys in Geophysics*, *37*(2), 269–305. https://doi.org/10.1007/s10712-016-9362-6

Crétaux, J.-F., Biancamaria, S., Arsen, A., Bergé-Nguyen, M., & Becker, M. (2015). Global surveys of reservoirs and lakes from satellites and regional application to the Syrdarya river basin. *Environmental Research Letters*, *10*(1), 15002. https://doi.org/10.1088/1748-9326/10/1/015002

Duan, Z., & Bastiaanssen, W. G. M. (2013). Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sensing of Environment*, *134*, 403–416. https://doi.org/10.1016/j.rse.2013.03.010

Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., et al. (2005). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews*, *81*(02), 163. https://doi.org/10.1017/s1464793105006950

El-Shazli, A., & Hoermann, G. (2016). Development of storage capacity and morphology of the Aswan High Dam Reservoir. *Hydrological Sciences Journal*, *61*(14), 2639–2648. https://doi.org/10.1080/02626667.2016.1151979

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, *45* RG2004. https://doi.org/10.1029/2005rg000183

Feyisa, G. L., Meilby, H., Fensholt, R., & Proud, S. R. (2014). Automated water extraction index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment*, *140*(October 2016), 23–35. https://doi.org/10.1016/j.rse.2013.08.029

Gao, H. (2015). Satellite remote sensing of large lakes and reservoirs: From elevation and area to storage. *Wiley Interdisciplinary Reviews: Water*, *2*(2), 147–157. https://doi.org/10.1002/wat2.1065

Gao, H., Birkett, C., & Lettenmaier, D. P. (2012). Global monitoring of large reservoir storage from satellite remote sensing. *Water Resources Research*, *48*, W09504. https://doi.org/10.1029/2012WR012063

Google Earth Engine. (n.d.). Simple cloud score. Retrieved August 6, 2019, from https://code.earthengine.google.com/5523bdba1566eb5809d8429a1dccfa34

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2016). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Heathcote, A. J., del Giorgio, P. A., Prairie, Y. T., & Brickman, D. (2015). Predicting bathymetric features of lakes from the topography of their surrounding landscape. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(5), 643–650. https://doi.org/10.1139/cjfas-2014-0392

Hollister, J., & Milstead, W. B. (2010). Using GIS to estimate lake volume from limited data. *Lake and Reservoir Management*, *26*(3), 194–199 . https://doi.org/10.1080/07438141.2010.504321

Jiang, L., Nielsen, K., Andersen, O. B., & Bauer-Gottwein, P. (2017). CryoSat-2 radar altimetry for monitoring freshwater resources of China. *Remote Sensing of Environment*, *200*, 125–139. Retrieved from. http://www.sciencedirect.com/science/article/pii/S0034425717303681, https://doi.org/10.1016/j.rse.2017.08.015

Joeckel, R. M., & Diffendal, R. F. (2004). Geomorphic and environmental change around a large, aging reservoir: Lake C. W. McConaughy, Western Nebraska, USA. *Environmental and Engineering Geoscience*, *10*(1), 69–90. https://doi.org/10.2113/10.1.69

Kansas Lakes and Reservoirs. (2016). Retrieved March 30, 2019, from https://kars.ku.edu/maps/kansaslakes/

Kress, W. H., Sebree, S. K., Littin, G. R., Drain, M. A., & Kling, M. E. (2005). Comparison of preconstruction and 2003 comparison of preconstruction and 2003 bathymetric and topographic surveys of Lake McConaughy, Nebraska. *U.S. Geological Survey Scientific Investigations Report 2005-5040*.

Lake History. (2016). Retrieved from https://ilovelakemac.com/lake-info/lake-history/

Liang, K., & Yan, G. (2017). Application of landsat imagery to investigate lake area variations and relict Gull Habitat in Hongjian Lake, Ordos Plateau, China. *Remote Sensing*, *9*(10). https://doi.org/10.3390/rs9101019

Lu, S., Wu, B., Yan, N., & Wang, H. (2011). Water body mapping method with HJ-1A/B satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, *13*(3), 428–434. https://doi.org/10.1016/j.jag.2010.09.006

McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, *17*(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*(1), 7. https://doi.org/10.1038/ncomms13603

Moradi, A., Metivier, L., de Viron, O., Calmant, S., & Mering, C. (2014). Evaluation of MODIS data for improved monitoring of the Caspian Sea. *International Journal of Remote Sensing*, *35*(16), 6060–6075. https://doi.org/10.1080/01431161.2014.943324

O'Loughlin, F. E., Neal, J., Yamazaki, D., & Bates, P. D. (2016). ICESat-derived inland water surface spot heights. *Water Resources Research.*, *52*(4), 3276–3284. https://doi.org/10.1002/2015WR018237

Pass locator: Aviso+. (n.d.). Retrieved February 27, 2019, from https://www.aviso.altimetry.fr/en/data/tools/pass-locator.html

Rahmani, V., Kastens, J. H., de Noyelles, F., Jakubauskas, M. E., Martinko, E. A., Huggins, D. H., et al. (2018). Examining storage capacity loss and sedimentation rate of large reservoirs in the Central U.S. great plains. *Water (Switzerland)*, *10*(2), 1–17. https://doi.org/10.3390/w10020190

Solander, K. C., Reager, J. T., & Famiglietti, J. S. (2016). How well will the Surface Water and Ocean Topography (SWOT) mission observe global reservoirs? *Water Resources Research*, *52*, 2123–2140. https://doi.org/10.1002/2015wr017952

Sulistioadi, Y. B., Tseng, K. H., Shum, C. K., Hidayat, H., Sumaryono, M., Suhardiman, A., et al. (2015). Satellite radar altimetry for monitoring small rivers and lakes in Indonesia. *Hydrology and Earth System Sciences*, *19*(1), 341–359. https://doi.org/10.5194/hess-19-341-2015

TOPEX/Poseidon Fact Sheet. (n.d.). Retrieved August 5, 2019, from https://sealevel.jpl.nasa.gov/missions/topex/topexfactsheet/

Tseng, K.-H., Shum, C. K., Kim, J.-W., Wang, X., Zhu, K., & Cheng, X. (2016). Integrating Landsat imageries and digital elevation models to infer water level change in Hoover Dam. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(4), 1696–1709. https://doi.org/10.1109/jstars.2015.2500599

USGS (2018). Landsat 4-7 Surface Reflectance (Ledaps) Product Guide. *Usgs.*, *1974*, 1–14. https://doi.org/10.1016/0042-207X(74)93024-3

Wang, X., Chen, Y., Song, L., Chen, X., Xie, H., & Liu, L. (2013). Analysis of lengths, water areas and volumes of the Three Gorges Reservoir at different water levels using Landsat images and SRTM DEM data. *Quaternary International*, *304*, 115–125. https://doi.org/10.1016/j.quaint.2013.03.041

Weekley, D. (2019a). Lake McConaughy Lake Dynamic Tracking Data. https://doi.org/10.6084/m9.figshare.8080664

Weekley, D. (2019b). Lake McConaughy Lake Dynamics Google Earth Engine Script. Retrieved from https://code.earthengine.google.com/8f3d01ac4acab9d5164c993f1d518270

Xiong, L., Deng, R., Li, J., Liu, X., Qin, Y., Liang, Y., & Liu, Y. (2018). Subpixel surface water extraction (SSWE) using Landsat 8 OLI data. *Water (Switzerland)*, *10*(5), 1–15. https://doi.org/10.3390/w10050653

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, *27*(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Yuan, L. T., Jung, H., Aierken, H. C., Beighley, A., Alsdorf, E., Tshimanga, R. M., & Kim, D. (2017). Absolute water storages in the Congo River floodplains from integration of InSAR and satellite radar altimetry. *Remote Sensing of Environment*, *201*, 57–72. https://doi.org/10.1016/j.rse.2017.09.003

Zhang, S., Foerster, S., Medeiros, P., de Araújo, J. C., Motagh, M., & Waske, B. (2016). Bathymetric survey of water reservoirs in north-eastern Brazil based on TanDEM-X satellite data. *Science of the Total Environment.*. https://doi.org/10.1016/j.scitotenv.2016.07.024
Zwally, D. H. J. (n.d.). NASA: ICESat. Retrieved August 5, 2019, from https://icesat.gsfc.nasa.gov/icesat/glas.php